REMARKS/ARGUMENTS

Favorable reconsideration of this application in light of the following discussion is respectfully requested.

Claims 1, 4, 7, 12, 17, 19, 21-31, 34, 35, 39, 43, 50-53, 56, 65, 70, 73-75, and 79-81 are pending in the present application. The claims are maintained by the present response.

In the outstanding Office Action, Claims 1, 4, 7, 12, 17, 19-31, 34-36, 39, 43, 50-53, 56, 65, 70, and 73-75 were rejected under 35 U.S.C. § 103(a) as unpatentable over Liddy et al. (U.S. Patent No. 5,873,056, herein "Liddy"), Kishi (U.S. Patent No. 6,041,329), and Wical (U.S. Patent No. 5,930,788), which is respectfully traversed.

Applicants thank Examiner Pardo for the courtesy of an interview extended to Applicant's representative on June 1, 2004. During the interview differences between the claims and the applied art were discussed. The Examiner indicated she would further review the pending claims in view of a filed response. Arguments presented during the interview are reiterated below.

Applicants thank the Examiner for returning the PTO-1449 form submitted with the previous Amendment. However, Applicants note that two references on that form were not initialed. Accordingly, Applicants respectfully request that all references on the PTO-1449 form be initialed as acknowledged. For the Examiner's convenience, a copy of the Information Disclosure Statement (IDS) provided with the previous Amendment is enclosed with this response.

Independent Claim 1 is directed to a computer processing apparatus that classifies a document and includes a database that has (i) a classification scheme, (ii) a classified vocabulary, and (iii) a classification data set. The classification scheme has a plurality of different subject matter categories, the classified vocabulary includes a plurality of terms in

each of the different subject matter categories, with each term being classified in accordance with the classification scheme, and the classification data set includes *a plurality of groups of terms* with each group being associated to a specific different one of the subject matter categories and each group including a plurality of terms exemplifying the associated category. Independent Claim 31 is a method claim that recites similar features as independent Claim 1.

In a non-limiting example, Figure 1 shows the classification scheme 20 having different subject matter categories 21. In another non-limiting example, Figure 3 shows the classified vocabulary 30, and the specification discloses at pages 61-62 the classification data set including a plurality of groups of terms (last paragraph on page 61 and first paragraph on page 62).

The outstanding Office Action recognizes at page 4, first full paragraph, that·

> neither Liddy nor Kishi teaches a classification data set comprising a plurality of groups of terms with each group being associated with a specific different one of the subject matter categories and each group including a plurality of terms exemplifying the associated category.

The outstanding Office Action relies on <u>Wical</u> for teaching those features. More specifically, the outstanding Office Action specifically states that a term "bank" disclosed by <u>Wical</u> at column 1, lines 44-62, belongs to a first category "finance & investment" and also to a second category "bodies of water," and these two categories correspond to the plurality of groups of terms recited in Claim 1.

However, as discussed during the interview, the term "bank" in <u>Wical</u> corresponds to the term "bayonet" disclosed in the specification, in Appendix A at page 61, and the two categories disclosed by <u>Wical</u> correspond to the "technology" and "photography" categories disclosed in the specification, also in Appendix A at page 61. Thus, the claimed plurality of groups is neither taught nor disclosed by <u>Wical</u>.

Further, Wical is directed to document *themes* or categories as disclosed at column 3, lines 19-28, and not to plurality of groups of terms as recited in independent Claim 1.

The themes of Wical represent a different concept from the groups of terms exemplifying categories for facilitating disambiguation between different meanings of a same term, as required in Claim 1. This is made particularly clear in Wical because the document themes may themselves be categories in the classification hierarchy, and the theme-based approach does not enable disambiguation between truly ambiguous terms.[1]

In addition, a critical requirement of Wical is that a thematic profile or a preliminary classification profile must be assigned to each document. Wical then uses an intermediate level of paragraph themes to identify the thematic content of the document as a whole. This reliance on a *level of paragraph thematisation*, critical to Wical, differs from Claim 1.

Moreover, the underlined theory of operation for the disambiguation process of Wical is that themes are not classified for a document if the themes are not verified. This is different from Claim 1, which does not require any such verification process. On the contrary, Claim 1 enables a document to be assigned to a category based solely on the extent in which the document includes a subset of terms from the group of terms exemplifying the category. The themes of Wical are topics or themes identified in the document, and the themes of Wical do not include a group of terms exemplifying a category, as discussed above and required by Claim 1.

Further, Applicants respectfully submit that there is no suggestion or motivation in either of the applied art to combine Liddy, Kishi, and Wical as suggested in the outstanding Office Action. Applicants note that to establish a proper basis for rejection based on a combination of references, MPEP § 2142 states:

---

[1] Wical, column 5, line 66, to column 6, line 11.

To establish a *prima facie* case of obviousness, three basic criteria must be met. First, there must be some suggestion or motivation, either in the references themselves or in the knowledge generally available to one of ordinary skill in the art, to modify the reference or to combine reference teachings. Second, there must be a reasonable expectation of success. Finally, the prior art reference (or references when combined) must teach or suggest all the claim limitations. The teaching or suggestion to make the claimed combination and the reasonable expectation of success must both be found in the prior art, and not based on applicant's disclosure.

A discussion of the operating principles of Liddy and Wical is considered to be in order here for showing that there is no suggestion or motivation in the applied art, as required by MPEP § 2142.

Liddy uses a lexical database such as a machine-readable type of Longman dictionary (LDOCE) to allocate subject codes to each word in a text to be processed. A representation of the meaning (context) of unformatted naturally occurring text is generated in a form of subject field codes. The subject field code (SFC) vectors of incoming documents are then matched to query SFC vectors enabling the documents to be ranked based on similarities.

Therefore, in Liddy, after dehypenation, stemming and functional word removal processes, the words in a document to be classified are looked-up in the lexical database and the subject code or codes for each word's tagged part of speech is used. The disambiguation process described in Liddy involves a heuristic order of processes. These processes involve first identifying unique or frequent subject codes, and then a computation is made to whether any subject code in a sentence equals or exceeds a predetermined frequency criterion. If the frequency criterion is exceeded or a subject code is the same as another subject code that was identified as a unique or frequent subject code, then a word is assigned that subject code.

However, if neither the frequency criterion nor a correspondence to a previously assigned unique or frequent subject code for the sentence is met, the system in Liddy provides for disambiguation via a corpus based on subject code correlation. This corpus is a

correlation matrix as shown in Table B of <u>Liddy</u>. The correlation matrix is obtained by correlating pairs of subject field codes in a corpus of text of the same type as that to be classified by the system. Thus, as can be seen from Table B in <u>Liddy</u>, the correlation matrix correlates pairs of subject field codes in accordance to a corpus of text of the same type as that to be classified.

Further, <u>Liddy</u> resolves one ambiguous word at a time by accessing the correlation matrix via the unique and high frequency subject codes determined for a sentence including the word. Then, the system evaluates the correlation coefficients between the unique frequent subject codes of the sentence and each of the multiple subject codes assigned to the word being disambiguated in order to determine which of the multiple subject codes has the highest correlation with the unique or high frequency subject codes. The system then selects the subject code as the unambiguous representation of the sense of the word, that is, the single subject code for the word.

Accordingly, <u>Liddy</u> describes a ***subject code correlation*** approach to disambiguation.

<u>Kishi</u> is simply concerned with an automated message processing system configured to automatically manage introduction of the movement of data storage media into a media library.

<u>Wical</u> classifies documents based on categories in a hierarchical knowledge catalog to generate a preliminary document classification profile as disclosed at column 3, lines 49-56. The disambiguation processing is effected based on the preliminary document classification thematic profiles and thematic profiles for the documents using the knowledge base and a category cross-reference database, which includes a list of category pairs.[2]

<u>Wical</u> uses the knowledge base to analyze hierarchical relationships among categories preliminarily classified for a document. For example, if the preliminary classifications

---

[2] <u>Wical</u>, column 4, lines 40-56, and column 4, lines 62-67.

include "OS/2" and "operating system," the disambiguation processing will validate "OS/2" as a type of operating system.

The disambiguation processing of <u>Wical</u> utilizes the category cross-reference database to validate a classification based on the category pairs. For example, <u>Wical</u> states at column 5, lines 13-19, that:

> the category "computer multimedia" is validated as a proper classification if the document also includes themes classified in the category "computer software industry" because the category pair "computer_multimedia & computer_software_industry" is located in the category cross-reference database.

The underlined theory of operation for the disambiguation system of <u>Wical</u> is that themes are not classified for a document if the themes are not verified. Thus, themes are not classified for a document in lieu of classifying themes in the wrong category. Consequently, categories included in the preliminary document classification profile do not appear in the verified document classification profile if the classification cannot be verified.

Therefore, in <u>Wical</u>, whether the preliminary classified category for a selected theme is valid is determined by analyzing the relationships between the category preliminary classified for the theme and other categories classified for different themes in the document. The disambiguation processing utilizes a category cross-references database, which includes a list of category cross-referenced pairs, to disambiguate categories assigned to themes by pairing a category classified for a theme and other categories classified for other themes in the document and by comparing these category pairs with category cross-reference database pairs. If a match occurs, then the categories of a document category pair are validated.

Thus, <u>Wical</u> discloses *a **theme-based*** approach to disambiguation, contrary to the ***subject code correlation*** approach disclosed in <u>Liddy</u>, which are different approaches.

In other words, Liddy does not generate the themes required by Wical and accordingly, one of ordinary skill in the art would not combine the theme-based disambiguation approach of Wical with the subject code correlation approach of Liddy.

Therefore, Applicants respectfully submit that there is no suggestion or motivation in the applied art to combine the different approaches of Liddy and Wical.

Further, there is no expectation provided in the references that such a modification of the approach in Liddy based on the approach in Wical would lead to any benefits of the disambiguation process, especially since the two processes are incompatible. Finally, it is respectfully submitted that no suggestion to modify the approach in Liddy based on the approach in Wical is found in the applied art. Although the outstanding Office Action states at page 4, third full paragraph, that combining Wical with the systems disclosed in Liddy and Kishi, would "increase the accuracy of retrieving documents from the database by disambiguating categories assigned to terms of keywords," Applicants respectfully submit that none of the applied art makes that suggestion.

Accordingly, it is respectfully submitted that independent Claims 1 and 31 and each of the claims depending therefrom patentably distinguish over the combination of Liddy, Kishi, and Wical.

With respect to independent Claim 12, as discussed in the previously filed amendment, independent Claim 12 is directed to a computer processing apparatus that includes a database that has a classified vocabulary including a plurality of terms in each of a plurality of different subject matter categories with each term being classified in accordance with the subject matter category structure of the database, and also includes a plurality of collocations, each associated with a specific different one of the subject matter categories and

each including a plurality of terms exemplifying the category associated for disambiguating different meanings of the same term.

However, as discussed above, none of the applied art teaches or suggests such features. Accordingly, it is respectfully submitted that independent Claim 12, independent Claim 39, which is a method claim corresponding to Claim 12, and each of the claims depending therefrom patentably distinguish over the applied art.

Independent Claim 53 requires a database having the features of the database of Claim 1, independent Claim 56 requires a database having the features of the database of Claim 12, and Claims 1 and 12 are believed to be allowable, as noted above. Accordingly, it is respectfully submitted that independent Claims 53 and 56 are also allowable.

Claim 65 is directed to an apparatus for classifying electronic documents and includes a storage unit for storing a classification scheme having a plurality of collocations in which each collocation is associated with a respective different subject matter area and includes a set of terms that exemplify that subject matter area. The apparatus also includes a unit for comparing terms used in the document to be classified with the terms in the collocations, a unit for allocating the document classified to one of the collocations that the comparing unit identifies as having the most number of terms in common with the document being classified, a unit for associating the document being classified to a code representing the subject matter area of the allocated collocation, and a unit for storing the document together with the associated code.

As discussed above, Liddy, Kishi, and Wical do not teach or suggest the use of a database that includes a classified vocabulary having a plurality of terms in each of a plurality of different subject matter categories with each term classified in accordance with the subject matter category structure of the database and also a plurality of collocations, each associated

with a specific different one of the subject matter categories and each including a plurality of terms exemplifying the associated category for disambiguating different meanings of the same term, as required in Claim 65.

Accordingly, it is respectfully submitted that independent Claim 65 and each of the claims depending therefrom patentably distinguish over the combination of Liddy, Kishi, and Wical.

Further, independent Claim 75 has been discussed in the previously filed amendment and Applicants respectfully submit that the combination of Liddy, Kishi, and Wical does not teach or suggest a processor readable medium storing processor readable instructions for causing a processor to compare terms appearing in a document with terms in a database, and a database structure having (i) a classification scheme having a plurality of different subject matter categories, (ii) a classified vocabulary having a plurality of terms in each of the different subject matter categories, and (iii) a classification data set having a plurality of groups of terms, each group being associated with a specific different one of the subject matter categories.

In addition, independent Claims 79 and 80 were also discussed in the previously filed amendment and none of the applied art teaches or suggests their features.

Accordingly, it is respectfully submitted that independent Claims 75, 79, and 80 patentably distinguish over the applied art.

Claim 81 is directed to an apparatus that includes a database that has a plurality of groups of terms with each group being associated with a specific different one of subject matter categories and each group including terms that may be used to describe a function, appearance or relationship with other objects of classified terms in that category or that may

be used in relation to terms in that category to facilitate disambiguation between different meanings of a same term.

None of the applied art teaches or suggests the above features. Accordingly, it is respectfully submitted that independent Claim 81 patentably distinguishes over the applied art.

Consequently, in light of the above discussion, the present application is believed to be in condition for allowance and an early and favorable action to that effect is respectfully requested.

Respectfully submitted,

OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT, P.C.

Customer Number
22850

Tel: (703) 413-3000
Fax: (703) 413 -2220
(OSMMN 08/03)
GJM/SS/RFF:smi

Gregory J. Maier
Registration No. 25,599
Surinder Sachar
Registration No. 34,423
Attorneys of Record

I:\ATTY\RFF\1S\105130001\105130001-REQ.DOC

11